Matteo Riondato

CONTACTS

Box 1910 Brown University
115 Waterman Street
Providence, RI 02912, USA

Voice: +1-401-654-3216
E-mail: matteo@cs.brown.edu
WWW: `http://cs.brown.edu/~matteo`

RESEARCH INTERESTS

Data analytics, graph mining, randomized algorithms, social network analysis, algorithms for distributed/parallel architectures for big data, privacy issues in data analysis, statistical learning theory.

EDUCATION

Brown University, Providence, Rhode Island, USA
Ph.D., Computer Science, *in progress since September 2009*. Expected completion: May 2014
- Doctoral Dissertation: *Randomized algorithms for big data analytics*
- Advisor: Prof. Eli Upfal

M.S., Computer Science, May 2010

Università di Padova, Padua, Italy
Laurea Magistrale (M.S.) 110/110 *cum laude*, Computer Engineering, July 2009
- Master Thesis: *Top-k frequent itemsets mining through sampling*
- Advisors: Prof. Andrea Pietracaprina, Prof. Eli Upfal, Fabio Vandin

Laurea (B.S.), Information Engineering, July 2007
- Final Project: *Algorithmical foundations of cryptography*
- Advisor: Prof. Andrea Pietracaprina

PUBLICATIONS

[10] M. Riondato and F. Vandin. Finding the true frequent itemsets. In *Proceedings of the 14th SIAM International Conference on Data Mining, SDM 2014, Philadelphia, PA, USA, April 24 – 26, 2014*, SIAM, 2014 (*to appear*).

[9] M. Riondato and E. Upfal. Efficient discovery of association rules and frequent itemsets through sampling with tight performance guarantees. ACM Transactions on Knowledge Discovery from Data (*to appear*).

[8] M. Riondato and E. M. Kornaropoulos. Fast approximation of betweenness centrality through sampling. In *Proceedings of the 7th International Conference on Web Search and Web Data Mining, WSDM 2014, New York, NY, USA, February 24 – 28, 2014*, ACM, 2014 (*to appear*).

[7] M. Riondato, J. A. DeBrabant, R. Fonseca, and E. Upfal. PARMA: A parallel randomized algorithm for association rules mining in MapReduce. In X.-w. Chen, G. Lebanon, H. Wang, and M. J. Zaki, editors, *Proceedings of the 21st ACM International Conference on Information and Knowledge Management, CIKM 2012, October 29 – November 02, 2012, Maui, HI, USA*, pages 85–94. ACM, 2012.

[6] M. Riondato and E. Upfal. Efficient discovery of association rules and frequent itemsets through sampling with tight performance guarantees. In P. A. Flach, T. De Bie, and N. Cristianini, editors, *Machine Learning and Knowledge Discovery in Databases*, volume 7523 of *Lecture Notes in Computer Science*, pages 25–41. Springer, 2012.

[5] A. Pietracaprina, G. Pucci, M. Riondato, F. Silvestri, and E. Upfal. Space-round tradeoffs for MapReduce computationseduce computations. In U. Banerjee, K. A. Gallivan, G. Bilardi, and M. Katevenis, editors, *International Conference on Supercomputing, ICS 2012, Venice, Italy, June 25–29, 2012*, pages 235–244. ACM, 2012.

[4] M. Akdere, U. Çetintemel, M. Riondato, E. Upfal, and S. B. Zdonik. Learning-based query performance modeling and prediction. In *Proceedings of the 28th International Conference on Data Engineering, ICDE 2012, April 1–5, 2012, Washington, DC, USA*, pages 390–401. IEEE Computer Society, 2012.

[3] M. Riondato, M. Akdere, U. Çetintemel, S. B. Zdonik, and E. Upfal. The VC-dimension of SQL queries and selectivity estimation through sampling. In D. Gunopulos, T. Hofmann, D. Malerba, and M. Vazirgiannis, editors, *Machine Learning and Knowledge Discovery in Databases - European Conference, ECML PKDD 2011, Athens, Greece, September 5–9, 2011, Proceedings, Part II*, volume 6912 of *Lecture Notes in Computer Science*, pages 661–676. Springer, 2011.

[2] M. Akdere, U. Çetintemel, M. Riondato, E. Upfal, and S. B. Zdonik. The case for predictive database systems: Opportunities and challenges. In *CIDR 2011, Fifth Biennial Conference on Innovative Data Systems Research, Asilomar, CA, USA, January 9–12, 2011, Online Proceedings*, pages 167–174. www.cidrdb.org, 2011.

[1] A. Pietracaprina, M. Riondato, E. Upfal, and F. Vandin. Mining top-k frequent itemsets through progressive sampling. *Data Mining and Knowledge Discovery*, 21(2):310–326, 2010.

UNDER SUBMISSION / IN PREPARATION

[11] A. Anagnastopoulos, L. Becchetti, A. Fazzone, I. Mele, and M. Riondato. The importance of being experts: Efficient max-finding in crowdsourcing.

[12] M. Riondato, D. García-Soriano, and F. Bonchi. Graph summarization with guarantees.

[13] M. Riondato, M. Akdere, U. Çetintemel, S. B. Zdonik, and E. Upfal. The VC-dimension of SQL queries and selectivity estimation through sampling (extended version).

AWARDS

Brown University Dissertation Fellowship **Fall 2013**

Yahoo! Research Summer Internship **Summer 2013**

Research Fellowship from MIUR of Italy under Project AlgoDeep prot. 2008TFBWL4 **Summer 2011**

Brown University Graduate Fellowship **Academic Year 2009–10**

INVITED TALKS

Fast betweenness estimation through sampling. *Data Management Group Seminar, Boston University*, 10/17/2013.

Fast betweenness estimation through sampling. *Lab Research Seminar, Yahoo! Labs Barcelona*, 6/13/2013.

Fast betweenness estimation through sampling. *Advanced Computing Group Talk, Department of Information Engineering, University of Padua, Italy*, 5/30/2013.

Statistical learning theory meets knowledge discovery: Randomized algorithms for Big Data analytics. *Brown CS Industrial Partners Program Symposium*, 4/25/2013.

Approximate aggregate database queries through sampling. *Invited Lecture for CSCI-2950-T, Brown University*, 10/25/2011.

Graphs algorithms in MapReduce: Design choices and optimizations. *Invited Lecture for CSCI-2950-U, Brown University*, 9/27/2011.

Statistical learning theory meets databases. *Advanced Computing Group Talk, Department of Information Engineering, University of Padua, Italy*, 06/24/2011.

Top-k frequent itemsets mining through sampling. *Advanced Computing Group Talk, Department of Information Engineering, University of Padua, Italy*, 9/16/2010.

ADDITIONAL RESEARCH EXPERIENCE

Yahoo! Labs, Barcelona, Spain
Summer Intern Research Scientist in the Web Mining Group **June – August 2013**

- Worked with Dr. Francesco Bonchi and others on algorithms for graph summarization and for frequent itemsets mining from data streams in a distributed setting.

Summer School on Massive Data Mining, Copenhagen, Denmark
Student **August 8 – 10, 2012**

- Attended the Summer School on Massive Data Mining organized by Prof. R. Pagh at IT University of Copenhagen.

Sapienza Università di Roma, Rome, Italy
Visiting Ph.D. Student **June – July 2012**

- Worked with Prof. S. Leonardi, A. Anagnostopoulos, and L. Becchetti on algorithms for a model of crowdsourcing computation, and on algorithms for MapReduce.

Università di Padova, Padua, Italy
Research Fellow **May – July 2011**

- Worked with Prof. A. Pietracaprina, G. Pucci, and F. Silvestri on "The MapReduce Paradigm: computational model and algorithms". Funded by MIUR of Italy under Project AlgoDEEP prot. 2008TFBWL4.

Chalmers University of Technology, Gothemburg, Sweden
Visiting Ph.D. Student **August 2010**

- Worked with Prof. Devdatt Dubhashi and helped organize a seminar on Probability and Computing.

Brown University, Providence, RI, USA
Visiting Grad Student **October 2008 – June 2009**

- Worked with Prof. Eli Upfal on master thesis *Top-K Frequent Itemsets Mining through Sampling*.

Teaching and Advising Training and Experience

The Harriet W. Sheridan Center for Teaching and Learning (Brown University), Providence, RI, USA

- Teaching Certificate I: Reflective Learning, AY 2013–14

Brown University, Providence, Rhode Island, USA

- Mentor, New Scientist Program Graduate-Undergraduate Mentoring Initiative, Spring 2014
- Teaching Assistant, Probability and Computing, Spring 2012, Spring 2013, Spring 2014
- Teaching Assistant, Probabilistic Methods in Computer Science, Fall 2010

Service

External/Sub- reviewer for the following conferences

- RANDOM'11, ICS'12, IPDPS'12, WSDM'13, MFCS'13, BigData'13, WSDM'14.

Brown University Graduate Student Council
President **April 2011 – December 2012**

- Represented the interests of the entire graduate student community (approx. 2000 students) with the university administration and the broader community at all levels. Previously held positions include Vice-President of Administration (Jan. – April 2011) and representative for the Computer Science Department (Sept. 2009 – April 2011, Jan 2013 – *ongoing*).

Brown University Graduate Council
Student Representative **September 2011 – May 2013**

- Represented the graduate student community in the highest body governing graduate education.

Brown University Strategic Planning Committee on Doctoral Education
Student Representative **September 2012 – May 2013**

- Represented the graduate student community to develop the presidential strategic plan for the next decade.

Brown Computer Science Theory Lunch
Organizer **January – December 2010**

- Responsible for organizing the weekly meeting of the Theory Group.

Membership

Society for Industrial and Applied Mathematics (SIAM)

- Student member and member of SIAM Activity Group on Data Mining and Analytics (SIAG/DMA).

Association for Computing Machinery (ACM)

- Student member and member of Special Interest Group on Knowledge Discovery from Data (SIGKDD).

Software

CentrSampl

- Algorithm to estimate node betweenness centrality in large graphs. Based on [8]. `http://cs.brown.edu/~matteo/centrsampl.tar.bz2` .

PARMA

- Frequent itemsets and association rules mining algorithm for Hadoop MapReduce. Based on [7]. `http://cs.brown.edu/~matteo/parma.tar.bz2` .

FreeSBIE
Developer, Release Engineer for the 2.x series **April 2004 – July 2009**

- Developed FreeSBIE, a Live-CD distribution of FreeBSD bootable from CD-ROM. Release Engineer for FreeSBIE 2.X. series, responsible for all aspects of the release. `http://www.FreeSBIE.org` .

The FreeBSD Project
src Committer **January 2006 – June 2013**

- Contributed to the development of the FreeBSD UNIX operating system. Granted write access to the main source repository. Worked on the `jail` security feature and on handling and solving bug reports of various nature. Author of the *Jail* chapter in the FreeBSD Handbook. `http://www.FreeBSD.org` .

References

Prof. Eli Upfal	Prof. Uğur Çetintemel	Dr. Francesco Bonchi
Dept. of Computer Science	Dept. of Computer Science	Yahoo! Labs Barcelona
Brown University	Brown University	Yahoo!
`eli@cs.brown.edu`	`ugur@cs.brown.edu`	`bonchi@yahoo-inc.com`

UP-TO-DATE VERSION Available from http://cs.brown.edu/~matteo/matteo_riondato_cv.pdf